

The Effect of Measurement Error on the Dose-Response Curve

by Isao Yoshimura*

In epidemiological studies for an environmental risk assessment, doses are often observed with errors. However, they have received little attention in data analysis. This paper studies the effect of measurement errors on the observed dose-response curve. Under the assumptions of the monotone likelihood ratio on errors and a monotone increasing dose-response curve, it is verified that the slope of the observed dose-response curve is likely to be gentler than the true one. The observed variance of responses are not so homogeneous as to be expected under models without errors. The estimation of parameters in a hockey-stick type dose-response curve with a threshold is considered on line of the maximum likelihood method for a functional relationship model. Numerical examples adaptable to the data in a 1986 study of the effect of air pollution that was conducted in Japan are also presented. The proposed model is proved to be suitable to the data in the example cited in this paper.

Introduction

In order to assess the risk of a chemical substance in the environment we have to estimate the dose-response relationship between the dose level of the substance and the prevalence rate of a set of symptoms that might be caused by it. The estimation is often performed based on the data in epidemiological studies.

In epidemiological studies, the raw data are often highly dispersed, as shown in Figure 1, which is a part of the data published by the Japan Environment Agency and is interpreted by Yoshimura (2). When a significant correlation is proved for such data, it is usually arranged in a reduced form (Fig. 2) by taking averages within categorized classes that are constructed on the dose. On such reduced figures we can easily confirm a monotone dose-response relationship.

However, there is one point to be noticed on this line of data processing. If the true dose-response relationship is similar to the one that is observed in Figure 2, the dispersion of raw data must be similar to that shown in Figure 3. In Figure 3 the middle solid line implies a dose-response curve, and the upper and the lower lines imply the widths of the standard deviations multiplied by 1.5 under the Poisson assumption stated later. The observed dispersion that is shown as open circles in Figure 3 is inhomogeneous in contrast with the expected dispersion shown by curves.

This paper gives a reasonable explanation of this inconsistency between the data and the fitted model by introducing a measurement error on doses and studies about the misleading effect of the measurement errors.

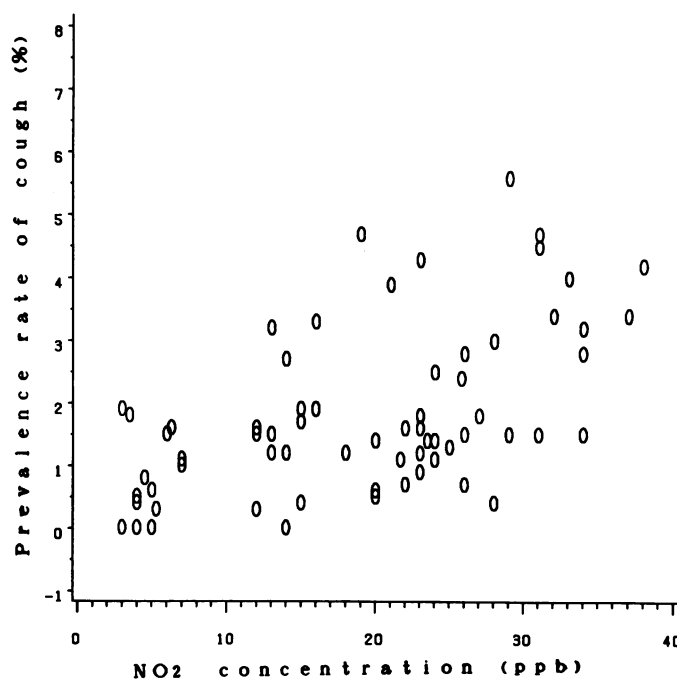


FIGURE 1. An example of epidemiological data (1). The dose variable is the NO_2 concentration (ppb) in the ambient air, and the response variable is the prevalence rate (%) of persistent cough among boys in each area. The number of areas in the study is 68.

Measurement Error Model

Consider that a areas are set in a epidemiological study and for each area A_i a dose variable and a response variable are observed. Let the observed variables be X_i and Y_i , $i = 1, 2, \dots, a$. Assume that X_i and Y_i are

*School of Engineering, Nagoya University, Nagoya 464-01, Japan.

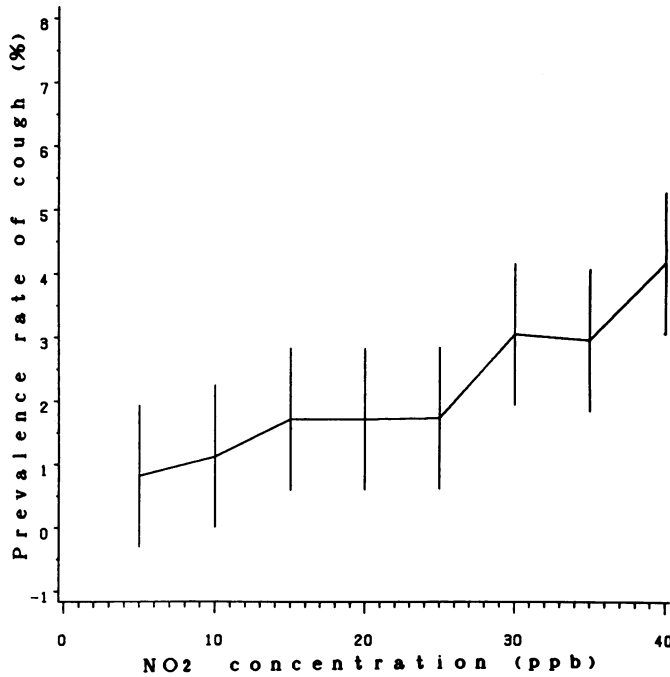


FIGURE 2. Observed average responses on categorized doses for data in Figure 1. The range of the estimated common standard deviation is indicated by vertical bars.

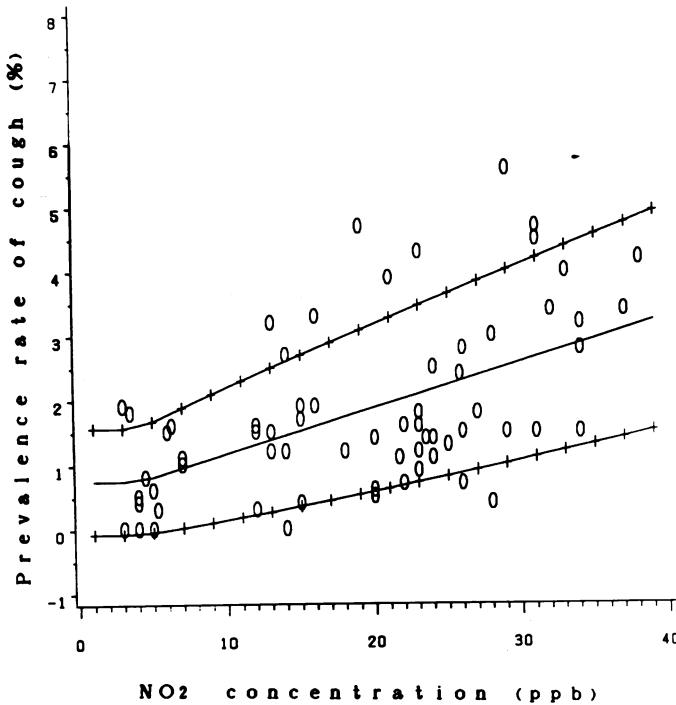


FIGURE 3. The dose-response curve estimated on the ignored measurement error. The line in the middle is the fitted dose-response curve: $\eta(\%) = 0.8 + 0.07(\xi(\text{ppb}) - 3.0)_+$, where $(\xi - a)_+ = \max(0, \xi - a)$. The modified two lines above and below the middle line imply the SD range multiplied by 1.5. The open circles are the observed values.

independently distributed random variables. In real situations the distributions may be discrete, but, for the convenience of expression, we regard that they are identified with probability densities $f(x; \xi_i, \sigma)$ and $g(y; \xi_i, \beta)$, respectively. It does not affect the following arguments.

When the areas are chosen purposively, ξ_i is regarded as an unknown parameter representing the true dose on the area A_i . On the contrary, when the areas are chosen randomly from a population of areas, ξ_i values are regarded as independently and identically distributed random variables with a prior probability density $h(\xi; \theta)$. Assume that the parameters β , σ , and θ are independent of areas. Let us eliminate the subscript i in the following, unless the specification of the area is necessary.

When the true dose ξ is given, the expectation of Y ,

$$\eta = E(Y | \xi) = \int y g(y; \xi, \beta) dy \quad (1)$$

is regarded as the true mean response on the area. As a function of ξ and β , $\eta = \eta(\xi; \beta)$ implies the true dose-response curve. In general, one of the principal purposes of such a survey is to know this true dose-response curve.

Effect of the Measurement Error

Functional Relationship Model

In the first case we assume ξ values are fixed unknown constants, which implies a functional relationship model. When we ignore measurement errors on the dose, we usually evaluate the average of the observed values of Y for the observed values of X as an estimate of the true dose-response curve. An example is shown in Figure 2. The dose-response curve thus obtained is regarded as an observation of an weighted average of $E(Y_i)$:

$$\begin{aligned} r_0(x; \beta, \sigma) &= \frac{\sum_{i=1}^a E(Y_i) \cdot f(x; \xi_i, \sigma)}{\sum_{k=1}^a f(x; \xi_k, \sigma)} \\ &= \sum_{i=1}^a E(Y_i) \cdot f^*(\xi_i; x, \sigma) \end{aligned} \quad (2)$$

where

$$f^*(\xi_i; x, \sigma) = \frac{f(x; \xi_i, \sigma)}{\sum_{k=1}^a f(x; \xi_k, \sigma)}. \quad (3)$$

This function $r_0(x; \beta, \sigma)$ is called the apparent dose-response curve. In many cases, conclusions derived from epidemiological studies are based on the apparent dose-response curve. However, if measurement errors exist, the apparent dose-response curve is distorted from the true one so as to cause a misunderstanding about the effect of the substance in question, as shown in the following material.

Consider the following assumptions on the distributions and parameters:

- Assumption 1: $\xi_1 \leq \xi_2 \leq \dots \leq \xi_a$ and at least one inequality is strict.

- Assumption 2: the region of x such that $f(x; \xi, \sigma) > 0$ makes an interval $[b_1, b_2]$, where b_1 and b_2 are independent of ξ and $b_1 < b_2$.
- Assumption 3: $r(\xi, \beta)$ is a strictly monotone increasing function of ξ .
- Assumption 4: The family of densities $f(x; \xi, \sigma)$ with the parameter ξ has a monotone likelihood ratio in x , that is, for any x^1, x^2, ξ^1 , and ξ^2 such that $x^1 < x^2$ and $\xi^1 < \xi^2$,

$$\frac{f(x^1; \xi^2, \sigma)}{f(x^1; \xi^1, \sigma)} < \frac{f(x^2; \xi^2, \sigma)}{f(x^2; \xi^1, \sigma)}. \quad (4)$$

Under these assumptions the following theorem holds:

Theorem 1. If assumptions 1 through 4 hold, then $r_0(x; \beta, \sigma)$ is a strictly monotone increasing function of x for $b_1 < x < b_2$ and constants c_1 and c_2 exist such that $\xi_1 < c_1 \leq c_2 < \xi_a$ and

$$r_0(x; \beta, \sigma) > r(x; \beta) \text{ for } b_1 < x < c_1 \quad (5)$$

$$r_0(x; \beta, \sigma) < r(x; \beta) \text{ for } c_2 < x < b_2. \quad (6)$$

Proof. By using Eq. (4) in Assumption 4 the following inequality holds for any x^1, x^2, ξ_i , and ξ_j such that $b_1 < x^1 < x^2 < b_2$ and $\xi_i < \xi_j$:

$$\begin{aligned} & \frac{f(x^1; \xi_j, \sigma) / \sum_{k=1}^a f(x^1; \xi_k, \sigma)}{f(x^1; \xi_i, \sigma) / \sum_{k=1}^a f(x^1; \xi_k, \sigma)} \\ & < \frac{f(x^2; \xi_j, \sigma) / \sum_{k=1}^a f(x^2; \xi_k, \sigma)}{f(x^2; \xi_i, \sigma) / \sum_{k=1}^a f(x^2; \xi_k, \sigma)}. \end{aligned} \quad (7)$$

Thus,

$$\frac{f^*(\xi_j; x^2, \sigma)}{f^*(\xi_i; x^1, \sigma)} < \frac{f^*(\xi_j; x^2, \sigma)}{f^*(\xi_j; x^1, \sigma)}. \quad (8)$$

This implies that as a family of distributions with parameter x , f^* has a monotone likelihood ratio in ξ . By the property of the family with monotone likelihood ratio [for example, see Lehmann (3)], the summation in Eq. (2) is strictly increasing in x . By Assumption 3, for any x such that $b_1 \leq x < \xi_1$,

$$\begin{aligned} r(x; \beta) & < r(\xi_1; \beta) \\ & = \sum_{i=1}^a r(\xi_i, \beta) \cdot f^*(\xi_i; x, \sigma) \\ & < \sum_{i=1}^a r(\xi_i, \beta) \cdot f^*(\xi_1; x, \sigma) \\ & = r_0(x; \beta, \sigma). \end{aligned} \quad (9)$$

Likewise, for $\xi_a < x < b_2$

$$r(x; \beta) > r_0(x; \beta, \sigma). \quad (10)$$

Since both $r(x; \beta)$ and $r_0(x; \beta, \sigma)$ are monotone increasing functions of x , the inequality of Eq. (5) holds for c_1 determined as $\inf(x)$ such that $r(x; \beta) \geq r_0(x; \beta, \sigma)$. Likewise Eq. (6) holds for c_2 determined as $\sup(x)$ such that $r(x; \beta) \leq r_0(x; \beta, \sigma)$.

This theorem implies that the apparent response tend to appear to be greater than the true one in low doses, while less in high doses. In most actual situations $c_1 = c_2$ as shown in the numerical examples in the next section, and then the average slope of the apparent dose-response curve is gentler than the true one in the central region of observed dose. Hence, it is defective as an estimate of the true dose-response curve and tends to cause an incorrect conclusion from the viewpoint of the risk assessment.

When both X and Y are normal variables and the dose-response curve is linear, this fact is well known in the context of the functional relationship model; however, in the situation we are faced with in epidemiological studies, the normality and the linearity are usually violated. I think this is the reason why this biased property contained in the apparent dose-response curve is neither noticed nor examined.

In the above theorem the strict monotonicity is assumed on the dose-response curve. If there is a threshold below which there is no increase of response, the monotonicity is not strict. Then Assumption 3 must be modified as Assumption 3' below:

- Assumption 3': $r(\xi; \beta)$ is constant for $\xi \leq d$ and is a strictly monotone increasing function of ξ for $d \leq \xi$, where the constant d is a given constant such that $\xi_1 < d < \xi_a$.

Even when Assumption 3 is replaced by Assumption 3' the above proof of the theorem is valid, so that the following corollary holds:

Corollary. If Assumptions 1, 2, 3', and 4 hold, then the conclusion of the theorem holds.

This corollary is important in actual situations, because even when there is a threshold it cannot be observed in data if there are errors in observed dose variables.

Consider next the effect of the errors in the doses on the variability of the response. The true variance of the response is reasonably defined as follows:

$$v(\xi; \beta) = V(Y) = \int \{y - r(\xi, \beta)\}^2 \cdot g(y; \xi, \sigma) dy \quad (11)$$

In contrast, the apparent variance can be defined as follows:

$$\begin{aligned} v_0(x; \beta, \sigma) & = \sum_{i=1}^a \{y - r_0(x; \beta, \sigma)\}^2 \\ & \quad \cdot g(y; \xi_i, \beta) \cdot dy \cdot f^*(\xi_i; x, \sigma) \\ & = \sum_{i=1}^a v(\xi_i; \beta) \cdot f^*(\xi_i; x, \sigma) \\ & \quad + \sum_{i=1}^a \{r(\xi_i; \beta) - r_0(x; \beta, \sigma)\}^2 \\ & \quad \cdot f^*(\xi_i; x, \sigma) \end{aligned} \quad (12)$$

If the true variance is constant on the whole range of the dose, the apparent variance is greater than it by the second term of Eq. (12). The excess variation in the apparent variance is in general, remarkable in a central part of the distribution of ξ values, as shown in the numerical example in the next section.

Structural Relationship Model

In the second model we assume ξ values are independently and identically distributed random variables with the prior probability density $h(\xi, \theta)$, which implies a structural relationship model. In this case, we modify the apparent dose-response curve as follows:

$$r_0(x; \beta, \sigma, \theta) = E(Y | X = x) \quad (13)$$

$$= \int r(\xi, \beta) g^*(\xi; x) d\xi,$$

where

$$g^*(\xi; x) = \frac{f(x; \xi, \sigma) \cdot h(\xi; \theta)}{\int f(x; \xi', \sigma) \cdot h(\xi'; \theta) d\xi'} \quad (14)$$

Assumptions are also modified as follows:

- Assumption 1': There are two constants a_1 and a_2 such that

$$a_1 < a_2, h(\xi; \theta) = 0 \text{ for } \xi \leq a_1 \text{ or } \xi \geq a_2. \quad (15)$$

- Assumption 2': The set of x such that $f(x; \xi, \sigma) > 0$ makes an interval $[b_1, b_2]$, where b_1 and b_2 are independent of ξ and $b_1 < a_1 < a_2 < b_2$.

Under these assumptions the following theorem holds:

Theorem 2. If Assumptions 1', 2', 3, and 4 hold, then $r_0(x; \beta, \sigma)$ is a strictly monotone increasing function of x for $b_1 < x < b_2$ and constants c_1 and c_2 exist such that $\xi_1 < c_1 \leq c_2 < \xi_a$ and

$$r_0(x; \beta, \sigma) > r(x; \beta) \quad \text{for } b_1 < x < c_1 \quad (16)$$

$$r_0(x; \beta, \sigma) < r(x; \beta) \quad \text{for } c_2 < x < b_2. \quad (17)$$

The proof is entirely the same as that of Theorem 1 and hence it is omitted. Likewise, when Assumption 3 is replaced with Assumption 3' the conclusion of the theorem holds.

As for the variance, the definition is modified as follows:

$$v_0(x; \beta, \sigma, \theta) = \int \int \{y - r_0(x; \beta, \sigma, \theta)\}^2 \cdot g(y; \xi, \beta) \cdot g^*(\xi; x) \cdot dy \cdot d\xi$$

$$= \int v(\xi, \beta) \cdot g^*(\xi; x) d\xi + \int \{r(\xi, \beta) - r_0(x; \beta, \sigma, \theta)\}^2 \cdot g^*(\xi; x) \cdot d\xi \quad (18)$$

Numerical Example

In this section, let us show some numerical results adaptable for the example shown in Figure 1. As for the dose, let X be such that $\ln(X)$ is distributed with $N(\ln(\xi), \sigma^2)$. As for the response, let Y be such that $n \cdot Y$ is distributed with Poisson($n\eta$), where n is supposed to be the number of persons sampled in the area and η is supposed to be the true prevalence rate in the area in the case of the example. Though Y is discrete in this model, it does not violate the validity of the argument in the preceding section.

Let the true dose-response curve be as follows:

$$\eta = r(\xi; \beta) = \beta_0 + \beta_1(\xi - \beta_2)_+, \quad (19)$$

where $(\xi - \beta_2)_+ = \max\{0, (\xi - \beta_2)\}$. This model implies the hockey-stick regression. In Eq. (19) β_0 is, in a sense, the spontaneous prevalence rate, β_1 is the risk factor, and β_2 is the threshold value.

Under the functional relationship model, the likelihood function L can be written as:

$$L = \text{const} \cdot \prod_{i=1}^a e^{(-n_i \eta_i)} (n_i \eta_i)^{n_i y_i} \quad (20)$$

$$\times (2\pi \sigma^2 x_i^2)^{-1/2} \exp \left\{ \frac{[\ln(x_i) - \ln(\xi_i)]^2}{-2\sigma^2} \right\}$$

where $\eta_i = r(\xi_i; \beta) = \beta_0 + \beta_1(\xi_i - \beta_2)_+$. If σ is known, the maximum likelihood estimates can be obtained numerically through an iterative method with suitable initial values of parameters. The estimates adaptable for the data shown in Figure 1 are obtained for some values of σ as shown in Table 1.

In order to check the adaptability of models, let us calculate the true and apparent dose-response curves for the following two cases:

Case 1: $\sigma = 0.4$, $\beta_0 = 0.008$, $\beta_1 = 0.0023/\text{ppb}$, $\beta_2 = 18 \text{ ppb}$, $\xi_i = x_i$.

Case 2: $\sigma = 0.4$, $\beta_0 = 0.008$, $\beta_1 = 0.0023/\text{ppb}$, $\beta_2 = 18 \text{ ppb}$, $n = 300$.

The prior distribution is

$$h(\xi; \theta) = 1/(\theta_2 - \theta_1) \quad \theta_1 < \xi < \theta_2, \quad (21)$$

where $\theta_1 = 5 \text{ ppb}$ and $\theta_2 = 35 \text{ ppb}$.

The one of them is a functional relationship model and the other is a structural relationship model. Both cases satisfy the assumptions in the corollary or its modified version.

For Case 1, the apparent dose-response curve can be obtained easily. For Case 2, some calculations are necessary and the result is as follows:

a) When $\theta_1 < \beta_2 < \theta_2$,

$$r_0(x; \beta, \sigma, \theta) = \beta_0 - \beta_1 \beta_2 \cdot \frac{\phi(u_2 - \sigma) - \phi(u_3 - \sigma)}{\phi(u_2 - \sigma) - \phi(u_1 - \sigma)}$$

$$+ \beta_1 \cdot x \cdot \exp(3\sigma^2/2) \cdot \frac{\phi(u_2 - 2\sigma) - \phi(u_3 - 2\sigma)}{\phi(u_2 - \sigma) - \phi(u_1 - \sigma)} \quad (22)$$

where $u_1 = (\ln\theta_1 - \ln(x))/\sigma$, $u_2 = (\ln\theta_2 - \ln(x))/\sigma$, $u_3 = (\ln\beta_2 - \ln(x))/\sigma$, and

Table 1. Estimates for various values of σ .

| σ | $\beta_0, \%$ | $\beta_1, \%/ \text{ppb}$ | β_2, ppb |
|----------|---------------|---------------------------|-----------------------|
| 0.2 | 0.9 | 0.23 | 19.0 |
| 0.3 | 0.8 | 0.23 | 18.0 |
| 0.4 | 0.8 | 0.23 | 18.0 |
| 0.5 | 0.8 | 0.23 | 16.0 |

$$\phi(u) = \int_{-\infty}^u \exp\{-v^2/2\} dv$$

b) When $\beta_2 \leq \theta_1$,

$$r_0(x; \beta, \sigma, \theta) = \beta_0 - \beta_1\beta_2 + \beta_1 \cdot x \cdot \exp(3\sigma^2/2) \cdot \frac{\phi(u_2 - 2\sigma) - \phi(u_1 - 2\sigma)}{\phi(u_2 - \sigma) - \phi(u_1 - \sigma)}. \quad (23)$$

Since the variance is the same as the mean for Poisson distribution, the true variance is given by

$$v(\xi; \beta) = \{\beta_0 + \beta_1(\xi - \beta_2)_+\}/n. \quad (24)$$

In contrast to it, the apparent variance is as follows:

a) When $\theta_1 < \beta_2 < \theta_2$,

$$\begin{aligned} v_0(x; \beta, \sigma, \theta) &= r_0(x; \beta, \sigma, \theta)/n - r_0^2(x; \beta, \sigma, \theta) \\ &+ \beta_0^2 + (\beta_1\beta_2 - 2\beta_0) \cdot \beta_1\beta_2 \\ &\cdot \frac{\phi(u_2 - \sigma) - \phi(u_3 - \sigma)}{\phi(u_2 - \sigma) - \phi(u_1 - \sigma)} \\ &+ 2(\beta_0 - \beta_1\beta_2)\beta_1 \cdot x \cdot \exp(3\sigma^2/2) \\ &\cdot \frac{\phi(u_2 - 2\sigma) - \phi(u_3 - 2\sigma)}{\phi(u_2 - \sigma) - \phi(u_1 - \sigma)} \\ &+ \beta_1^2 \cdot x^2 \cdot \exp(4\sigma^2) \\ &\cdot \frac{\phi(u_2 - 3\sigma) - \phi(u_3 - 3\sigma)}{\phi(u_2 - \sigma) - \phi(u_1 - \sigma)}. \end{aligned} \quad (25)$$

b) When $\beta_2 \leq \theta_1$,

$$\begin{aligned} v_0(x; \beta, \sigma, \theta) &= r_0(x; \beta, \sigma, \theta)/n - r_0^2(x; \beta, \sigma, \theta) \\ &+ (\beta_0^2 - \beta_1\beta_2)^2 + 2(\beta_0 - \beta_1\beta_2)\beta_1 \\ &\cdot x \cdot \exp(3\sigma^2/2) \\ &\cdot \frac{\phi(u_2 - 2\sigma) - \phi(u_1 - 2\sigma)}{\phi(u_2 - \sigma) - \phi(u_1 - \sigma)} + \beta_1^2 \\ &\cdot x^2 \cdot \exp(4\sigma^2) \\ &\cdot \frac{\phi(u_2 - 3\sigma) - \phi(u_1 - 3\sigma)}{\phi(u_2 - \sigma) - \phi(u_1 - \sigma)}. \end{aligned} \quad (26)$$

The result of the numerical calculation for Case 1 is shown in Figure 4. The result for Case 2 gives entirely the same figure as Figure 4, and so the figure is omitted. In the figure the hockey-stick type solid line implies the true dose-response curve, and the curved solid line implies the apparent dose-response curve calculated for $\sigma = 0.4$ and $\xi_i = x_i$. The two thick lines above and below the solid lines imply the width of the standard deviation multiplied by 1.5. Open circles are the observed points shown in Figure 1.

As far as the data shown in Figure 1 is concerned, the measurement error models with the parameters set in Case 1 and Case 2 are well fitted, compared with the

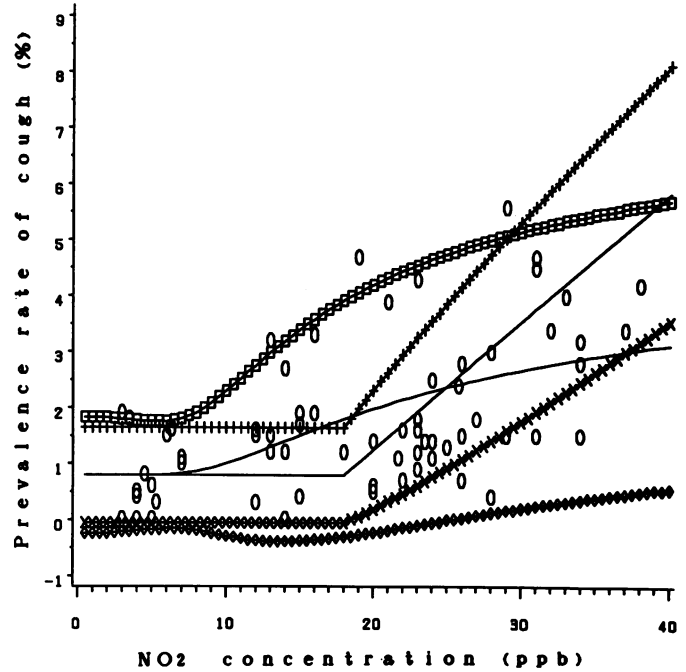


FIGURE 4. The dose-response curve estimated on a measurement error model. (a) The true dose-response curve: $\eta(\%) = 0.8 + 0.23(\xi(\text{ppb}) - 18.0)_+$, where $(\xi - a)_+ = \max(0, \xi - a)$. (b) The apparent dose-response curve calculated for $\sigma = 0.4$ and $\xi_i = x_i$. (c,d) The range of the standard deviation multiplied by 1.5. (e,f) The range of apparent SD multiplied by 1.5. The open circles are the observed values.

model without measurement error that is shown in Figure 3.

Conclusion and Discussion

It has been said that it is difficult to fit a simple dose-response curve to such data as that shown in Figure 1. However, in this paper the possibility of fitting a simple dose-response curve to such data is shown by assuming the existence of a measurement error on the dose. It is to be noted that if we ignore the measurement error—in spite of the actual existence of it—we are likely to estimate the true dose-response curve incorrectly with a bias.

Further investigations should obtain effective methods of estimation of parameters well fitted to real data under the measurement model. However, it is anticipated that the knowledge about the dispersion or the standard deviations is necessary to estimate parameters, because even when the normality of errors and the linearity of the curve are assumed, the knowledge about the dispersion inevitably get satisfactory result, as is noted in Fuller (4) or Singh and Kanji (5).

REFERENCES

1. Japan Environment Agency. A Survey of Respiratory Diseases Using Questionnaire [in Japanese]. JEA, Tokyo, Japan, 1986.

2. Yoshimura, I. On the determination of the regulatory criteria for air pollution. Paper presented at 46th session of the International Statistical Institute, Tokyo, Japan, invited paper 31.1, 1987.
3. Lehmann, E. L. Testing Statistical Hypotheses, second ed. John Wiley and Sons, New York, 1986.
4. Fuller, W. A. Measurement Error Models. John Wiley and Sons, New York, 1987.
5. Singh, M, and Kanji, G. K. Fitting a non-linear model with errors in both variables and its application. J. Appl. Stat. 15: 267–174 (1988).